Syllabus

Course Information

This course provides a hands-on introduction to using data to rigorously answer research questions. Students practice cleaning and manipulating data, creating data visualizations, and conducting introductory level statistical analysis using real-world data sets that are relevant to their field. Analysis topics include single and two-sample inference, analysis of variance, multiple regression, analysis of co-variance, experimental design, repeated measures, nonparametric procedures, and categorical data analysis. Reproducible research is strongly emphasized through the use of statistical computing software (e.g. SPSS, Stata, SAS, R, Python). Recommended for all majors that use data for research. 3 hours discussion.

Instructor

- **Name & pronouns:** Dr. Robin Donatello (Dr. D, she/her)
- **Website:** https://www.norcalbiostat.com/
- **Office Location:** Holt 202
- **E-mail:** rdonatello@csuchico.edu
- **Best Contact Method:** Discord
- **Student Office Hours:** TBD

Meeting Logistics

- **Meeting Pattern:** MW 4-5:15, Holt 155
- **Prerequisites:** Basic computer literacy. Recent statistics course such as Math 105, MATH 315, or MATH 350.
- **Mode of Instruction:** In Person.

Required materials and accounts

Homework 0 provides a checklist for you to make sure you have everything.

- **Class website**: https://norcalbiostat.github.io/MATH615/
- **Textbook (Required)**: *Practical Multivariate Analysis (PMA6), 6th ed* by Afifi, May, Donatello, Clark. [Link] There are several available in the library and I have two I can check out for the semester.
- **Textbook (Required)**: Introduction to Modern Statistics (IMS). Free e-book/pdf at https://openintro-ims.netlify.app
- **Laptop & Internet**: Expect to bring your laptop to class every day. Contact me if this poses a problem or concern for you. Refer to ITSS for help with internet.
- **Data Analysis software**: R and R Studio
- **Canvas**: Assignmnent submission, gradebook, class schedule/calendar
- **Google Drive**: Quizzes and collaborative work.
- **Discord**: A free discussion platform with a lot of collaborative functionality. The defacto method of communication for the class.

What are we going to learn? (Course Learning Outcomes)

By the end of the course, students will be able to...

- Import data into a statistical analysis software program in a format that is ready to analyze.
- Process, screen, recode, transform, and clean data.
- Describe distributions and patterns of data using visualizations and words.
- Select and carry out an appropriate statistical analysis.
- Explain study results and limitations to a non-technical audience.
- Understand and implement a reproducible research pipeline.
- Become a data nerd (Optional, but recommended).

Schedule of Topics

The general ordering of topics is:

- Data Collection and recording
- Preparing data for analysis
- Data Visualization

- Foundations for Inference: probability distributions, point and interval estimation, Hypothesis Testing
- Inference comparing multiple samples (t-tests, ANOVA, χ^2 tests)
- Study design, confounding, causation
- Linear regression analysis (Simple and Multiple, Categorical predictors, variable selection)
- Logistic regression analysis
- Model building techniques and comparing model fit

Pre-requisite statistics Knowledge

This class requires a recent statistics course such as Math 105, Math 315 or Math 350. Sometimes you need a refresher on the content that was covered in those classes (such as how to summarize data, probability, distributions). While we will be going over all the basic visualizations and tests again, it will be fast and the emphasis is on interpretation and proper use in research.

- Chapter 6 in the PMA6 textbook contains an overview on how to select appropriate summary measures, graphs (also in chapter 4), and analysis methods. (we'll go through this chapter)
- The Intro to Modern Statistics book can help with this (this is a free required textbook)
- I also like the OpenIntro Statistics book (also available as a free pdf) chapters 1-4.

Items that you need to know very well are

- types of data (categorical, quantitative, numerical, qualitative)
- appropriate plot types depending on types of data (histogram vs bar chart)
- measures of center (mean, median, mode)
- measures of spread (standard deviation, standard error, IQR, range)

Items that you need to know at least a little about

- rules of basic probability (e.g. is between 0 and 1, disjoint/mutually exclusive)
- the normal distribution: Z-scores (how to calculate, how to interpret) and probability under the curve (one-tail, two tail)

How do you get help with stuff?

- See the help page on the course website.

There is a lot of info on this page as I am trying to keep the syllabus reasonably short.

Coursework and Assignments

What am I graded on?

I care both that you are keeping up with the material, but also that by the end of the term you can conduct your own personalized analysis on a new data set. Here is how I check in on your progress throughout the semester

- **Quizzes** on the current or prior content. Quizzes are a great way to test yourself, enhance recall, and self-assess your comprehension of the subject. Quizzes are administered in Google Forms, and done in two parts.
  - **Individual**: First try to answer as many questions without looking at your notes. Then use your notes to answer the rest.
  - **Group**: After the individual quiz closes I will choose 1-4 questions that scored the lowest. We will do a class-wide discussion of those topics, then you will work in groups to redo the selected questions for half credit.
- **Assignments** This is your testing ground, your first round of practice. Can you take what you learned and apply it to a data set. Graded on completeness and effort.
- **Peer reviews** Helping others enhances your own understanding.
- **Project** Can you weave together data and story to do research? This is your second round of practice. Your work should be more polished than in your homework, and I will be grading them in more detail and using a rubric.
- **Comprehensive Final Exam** Can you identify and apply correct statistical theory to new situations?

Okay, but what about the points?

- The gradebook in Canvas contains columns for all graded materials with point values and rubrics.
- The project uses a *mastery based* grading system that allows you to continue to revise your work until it reaches an acceptable level of quality. The details are described on the Project page.
- See here for info on grading peer reviews.

Your final grade will be displayed as a running total in Canvas. The grades are weighted by category as follows:

- Assignments: 25%
- Quizzes: 25%
- Peer Reviews: 5%

- Comprehensive Exam: 15%
- Project: 30%

I use a standard grade cutoff of 100-90%: A, 89-80%: B, 79-70%: C.

Class Policies and statements

Everyone is welcome here

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity: gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture. Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

I would like to create a learning environment that supports a diversity of thoughts, perspectives and experiences, and honors your identities (including race, gender, class, sexuality, religion, ability, etc.) To help accomplish this:

- Let me know if you have a name and/or set of pronouns that differ from those that appear in your official Chico records. I make it a point to call on people by name, so please make sure that I know what you want to be called. - I also want to try to pronounce your name as accurately as possible. The more you help and correct me the better I can do to honor your name.
- If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. I want to be a resource for you. Remember that you can also submit anonymous feedback using the feedback button on the schedule and help pages of the website (which will lead to me making a general announcement to the class, if necessary to address your concerns).
- If you prefer to speak with someone outside of the course, the Office of Diversity and Inclusion is here to assist. Their number is 530-898-4764, and email diversityoffice@csuchico.edu
- I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to me about it. (Again, anonymous feedback is always an option).

Adapted from Monica Linden at Brown University.

Furthermore, I would like to acknowledge that Chico State stands on lands that were originally occupied by the first people of this area, the Mechoopda. I recognize their distinctive spiritual relationship with this land and the waters that run through campus. I am humbled that our campus resides upon sacred lands that once sustained the Mechoopda people for centuries.

Attendance

Class attendance is expected. Talk to me ahead of time if you need to miss a class for a planned reason. In the event of an unplanned reason, PM me in Discord when you can so that I know you are still alive.

- If you are not feeling well or are experiencing Covid symptoms **DO NOT COME TO CLASS**. DM me in Discord we'll make a backup plan.

Online class

This is a graduate class and you all are adults with lives outside this class. Things happen. Each class session will be live streamed, with the recording posted to Canvas within a few days. Common reasons to join the class virtually: - you're sick - you're out of town

Don't expect as good of quality of presentation, and this is not a long term solution/resource. Don't abuse this resource.

Late work

I don't penalize for late work, but I prioritize on time submissions. If you submit an assignment after I'm done grading everyone elses, and we're moving on to the next topic then your assignment may have to wait a week before I can get back to it. That is likely to negatively impact your ability to move forward and to get feedback in time.

There are exceptions for when your work is peer reviewed. There is no grace period for peer reviews. You must be responsible and timely for your colleagues.

Use of AI

As an instructor I recognize there are a variety of AI programs available to assist in creating text and writing code. However, I want to stress that, AI programs are not a replacement for human creativity, originality, and critical thinking. Writing (text and code) is a skill that you must nurture over time in order to develop your own individual voice, style, and view.

The use of chat GPT is allowed/encouraged **to help you learn how to code** but all code used must be fully explained in text. We will cover what this means and how to do this when we start data management around week 3. How ChatGPT can help you write code

You are responsible for fact checking the accuracy of statements composed by AI language models. These models are known to produce bullshit responses. And yes, this is a technical term.

> ❗ Be up front!
>
> When presenting text written by Artificial intelligence (AI) language models, such as ChatGPT, you must include an appropriate citation. https://apastyle.apa.org/blog/how-to-cite-chatgpt
> This also goes for when you have it write code for you. You need to indicate when and where you used it.

Collaboration

Collaboration on assignments is encouraged. People learn better when they have someone else to talk through concepts with. However your submitted work must represent **your personal effort** on all parts.

Not allowed

- Working with or getting help from others on exams and individual quizzes
- Copying code from another student's homework and presenting it as your own work.
- Copy/paste from AI tools or internet sources without customization, citation or explanation
- Getting your sibling/friend/colleague to write code for you
- Submitting any assignment that is not your own personal effort.

Allowed

- Helping each other solve homework problems (concepts or code)
- Use AI to help explain a concept
- Use AI to generate starter code that you modify for your own example.
- Use AI to write code that you cite/disclose and you explain in *your words* what it is doing in details.
- Copy/paste code from my course notes (this is actually encouraged!)

If at any time I suspect that the work you are submitting is not reflective of your personal knowledge I may ask you to verbally explain a piece of code to me. If your explanation is insufficient may result in a 0 on that assignment.

Any use outside of this permission constitutes a violation of Chico State's Integrity Policy and may result in you being reported to the Office of Students Rights and Responsibilities.

Campus Resources

Also check out the resources on the Help page.

Americans with Disabilities Act

If you need course adaptations or accommodations because of a disability or chronic illness, or if you need to make special arrangements in case the building must be evacuated, please make an appointment with me as soon as possible, or see me during office hours. Please also contact Accessibility Resource Center (ARC) as they are the designated department responsible for approving and coordinating reasonable accommodations and services for students with disabilities. ARC will help you understand your rights and responsibilities under the Americans with Disabilities Act and provide you further assistance with requesting and arranging

accommodations. We try our best to ensure equal access to materials in accessible formats. Reach out to your me if there is some aspect of the course materials that are inaccessible to you.

Accessibility Resource Center 530-898-5959 Student Services Center 170 arcdept@csuchico.edu http://www.csuchico.edu/arc

---

University Policies

Adding and Dropping the course

You are responsible for understanding the policies and procedures about add/drops, academic renewal, etc., found in the CSU Chico University Catalog. You should be aware of the new deadlines and penalties for adding and dropping classes.

Confidentiality and Mandatory Reoprting

As an instructor, one of my responsibilities is to help create a safe learning environment on our campus. I also have a mandatory reporting responsibility related to my role as a your instructor. I am required to share information regarding sexual misconduct with the University. Students may speak to someone confidentially by contacting the Counseling and Wellness Center (898-6345) or Safe Place (898-3030). Information on campus reporting obligations and other Title IX related resources are available here: www.csuchico.edu/title-ix.

Academic Integrity

Academic integrity is defined as "a commitment, even in the face of adversity, to five fundamental values: honesty, trust, fairness, respect, and responsibility". From these values flow principles of behavior that enable academic communities to translate ideals to action. Academic integrity is expected and required. No forms of cheating or plagiarism will be tolerated. Please see your student handbook at https://www.csuchico.edu/scrr/integrity.shtml if you have questions about the meaning of these terms or the consequences of violating academic integrity.